

RECEIVED

JUL 17 2001

TECH CENTER 1600/2900

SEQUENCE LISTING

A1

<110> NISHIMURA, SATORU
KOIKE, AYUMI



<120> CHOLINE MONOOXYGENASE GENE

<130> 0213-1431-0

<140> 09/534,995

<141> 2000-03-27

<150> JP 273275/1999

<151> 1999-09-27

<160> 17

<170> PatentIn version 3.1

<210> 1

<211> 1828

<212> DNA

<213> Chenopodium album

<220>

<221> CDS

<222> (129)..(1427)

<223>

<400> 1

gaactataca agctaagtta agcttaagct atattgttga tcattctttca tactacttcc	60
tttaaaaaaaaa aaattataac aacaaaagga agtgtgaatt ttttccttga tcattcatata	120
acatcaat atg gca gca agt gca aca aca atg ttg ctg aaa tac cca aca	170
Met Ala Ala Ser Ala Thr Thr Met Leu Leu Lys Tyr Pro Thr	
1 5 10	
act gta tgt ggt ata cca aat tca tca tca aac aat gat act tca aat	218
Thr Val Cys Gly Ile Pro Asn Ser Ser Ser Asn Asn Asp Thr Ser Asn	
15 20 25 30	
aat ata gtc cca att cca caa act agt act aat aat ccg gta ctt aag	266
Asn Ile Val Pro Ile Pro Gln Thr Ser Thr Asn Asn Pro Val Leu Lys	
35 40 45	
ttt cgt acc cct aat aaa acc att aac gcc gtc gct gcc ccg gct ttt	314
Phe Arg Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala Phe	
50 55 60	
cct tct tta aac acc acc act act ccg ccg tcg att caa tca ctt gtc	362
Pro Ser Leu Asn Thr Thr Thr Thr Pro Pro Ser Ile Gln Ser Leu Val	
65 70 75	
cag gaa ttc gat ccg aag att ccg gct aag gat gct ctt acg cct cct	410
Gln Glu Phe Asp Pro Lys Ile Pro Ala Lys Asp Ala Leu Thr Pro Pro	
80 85 90	
agc tct tgg tat act gac gct gct ttc tat gct cat gaa ctt gac cgt	458
Ser Ser Trp Tyr Thr Asp Ala Ala Phe Tyr Ala His Glu Leu Asp Arg	
95 100 105 110	
atc ttt tat aag gga tgg caa gtc cca ggg tac agt gat caa att aag	506
Ile Phe Tyr Lys Gly Trp Gln Val Pro Gly Tyr Ser Asp Gln Ile Lys	
115 120 125	
gag cct aac caa tat ttc acc gga acg tta gga aat gtt gaa tat ttg	554
Glu Pro Asn Gln Tyr Phe Thr Gly Thr Leu Gly Asn Val Glu Tyr Leu	
130 135 140	
gtg tgt cga gat ggt gaa gga aaa gtt cat gca ttt cac aac gtt tgc	602
Val Cys Arg Asp Gly Glu Gly Lys Val His Ala Phe His Asn Val Cys	
145 150 155	
acc cat cgt gct tcg att ctt gct tgt gga agt gga aaa aaa tcg tgt	650

Thr	His	Arg	Ala	Ser	Ile	Leu	Ala	Cys	Gly	Ser	Gly	Lys	Lys	Ser	Cys	
160						165					170					
ttt	gtg	tgc	cct	tac	cat	gga	tgg	gta	ttt	ggc	atg	aat	gga	tcg	ctt	698
Phe	Val	Cys	Pro	Tyr	His	Gly	Trp	Val	Phe	Gly	Met	Asn	Gly	Ser	Leu	
175					180					185					190	
aca	aaa	gct	tcc	aaa	gca	acc	gaa	gaa	cag	tca	ctt	gat	ccc	gat	gaa	746
Thr	Lys	Ala	Ser	Lys	Ala	Thr	Glu	Glu	Gln	Ser	Leu	Asp	Pro	Asp	Glu	
				195					200					205		
ctt	ggg	ctt	gta	ccc	ctg	aaa	gtt	gca	gta	tgg	ggc	cca	ttt	ata	ctc	794
Leu	Gly	Leu	Val	Pro	Leu	Lys	Val	Ala	Val	Trp	Gly	Pro	Phe	Ile	Leu	
			210					215					220			
ata	agt	ttg	gac	aga	tca	agc	ctt	gaa	gta	ggt	gat	gtt	gga	tct	gaa	842
Ile	Ser	Leu	Asp	Arg	Ser	Ser	Leu	Glu	Val	Gly	Asp	Val	Gly	Ser	Glu	
		225					230					235				
tgg	ctt	ggt	agt	tgt	gct	gaa	gat	gtt	aag	gcc	cat	gct	ttt	gac	cct	890
Trp	Leu	Gly	Ser	Cys	Ala	Glu	Asp	Val	Lys	Ala	His	Ala	Phe	Asp	Pro	
	240					245					250					
aat	tta	cag	ttc	atc	aat	agg	agt	gaa	ttt	cca	atg	gaa	tct	aat	tgg	938
Asn	Leu	Gln	Phe	Ile	Asn	Arg	Ser	Glu	Phe	Pro	Met	Glu	Ser	Asn	Trp	
255					260					265					270	
aag	att	ttc	agt	gac	aac	tat	ttg	gat	agc	tcg	tac	cat	gtt	cct	tat	986
Lys	Ile	Phe	Ser	Asp	Asn	Tyr	Leu	Asp	Ser	Ser	Tyr	His	Val	Pro	Tyr	
				275					280					285		
gca	cac	aag	tac	tat	gct	act	gaa	ctc	gac	ttt	gat	act	tac	caa	act	1034
Ala	His	Lys	Tyr	Tyr	Ala	Thr	Glu	Leu	Asp	Phe	Asp	Thr	Tyr	Gln	Thr	
			290					295					300			
gat	atg	atc	gga	aac	gtc	acg	att	caa	aga	gtg	gca	ggg	agt	tca	aac	1082
Asp	Met	Ile	Gly	Asn	Val	Thr	Ile	Gln	Arg	Val	Ala	Gly	Ser	Ser	Asn	
		305					310					315				
aat	ggt	ttt	aat	aga	ctt	gga	tct	caa	gca	ttc	tat	gct	ttt	gca	tac	1130
Asn	Gly	Phe	Asn	Arg	Leu	Gly	Ser	Gln	Ala	Phe	Tyr	Ala	Phe	Ala	Tyr	
	320					325					330					
cct	aac	ttt	gct	gtg	gaa	agg	tat	ggc	cct	tgg	atg	aca	aca	atg	cac	1178
Pro	Asn	Phe	Ala	Val	Glu	Arg	Tyr	Gly	Pro	Trp	Met	Thr	Thr	Met	His	
335					340					345					350	
att	ctt	cca	tta	gga	cca	agg	aaa	tgc	aaa	tta	gtg	gtg	gac	tac	tac	1226
Ile	Leu	Pro	Leu	Gly	Pro	Arg	Lys	Cys	Lys	Leu	Val	Val	Asp	Tyr	Tyr	
				355					360					365		
att	gaa	aaa	tca	aag	ctg	gac	gac	aag	gat	tac	atc	gag	aag	ggc	att	1274

Ile	Glu	Lys	Ser	Lys	Leu	Asp	Asp	Lys	Asp	Tyr	Ile	Glu	Lys	Gly	Ile		
			370					375					380				
gca	atc	aat	gat	aat	gta	cag	aaa	gaa	gat	gtg	gtg	ttg	tgt	gaa	agt	1322	
Ala	Ile	Asn	Asp	Asn	Val	Gln	Lys	Glu	Asp	Val	Val	Leu	Cys	Glu	Ser		
		385					390					395					
gtc	caa	aaa	ggg	ttg	gaa	aca	cca	gca	tat	cgt	agt	gga	aga	tat	gtg	1370	
Val	Gln	Lys	Gly	Leu	Glu	Thr	Pro	Ala	Tyr	Arg	Ser	Gly	Arg	Tyr	Val		
	400					405					410						
atg	cca	att	gag	aaa	gga	atc	cat	cat	ttc	cac	tgc	tgg	ttg	cac	caa	1418	
Met	Pro	Ile	Glu	Lys	Gly	Ile	His	His	Phe	His	Cys	Trp	Leu	His	Gln		
415					420				425						430		
gta	ttg	aag	tgattgcagc	agatcatcag	atgttcgttt	cttcttgtat										1467	
Val	Leu	Lys															
tggaattgga	tattatgatt	aataagtaaa	attataatgt	cataatgtag	ttgagattgt											1527	
tgctagagtt	gagcgtatgc	tcctcatgca	ctacttagtt	atcaagtgtg	tatgtctttg											1587	
gtcatgggca	aaatgtatgt	ttcttgctag	aatttatata	ttatgggtgct	aatgtccaat											1647	
ataaataaaa	accatagcac	ccctttaatt	ccctacttag	gtttatatcc	cattttatttt											1707	
cgggggatct	atgagataga	ttgtctatga	acattatttt	tcgactcgtg	tatgggtattc											1767	
atcccttggtg	taggggtgaag	taaacattga	gtgtatgaag	ttttcattga	gtttctgctt											1827	
t																1828	

<210> 2

<211> 433

<212> PRT

<213> Chenopodium album

<400> 2

Met	Ala	Ala	Ser	Ala	Thr	Thr	Met	Leu	Leu	Lys	Tyr	Pro	Thr	Thr	Val
1				5					10					15	

Cys	Gly	Ile	Pro	Asn	Ser	Ser	Ser	Asn	Asn	Asp	Thr	Ser	Asn	Asn	Ile
			20					25					30		

Val Pro Ile Pro Gln Thr Ser Thr Asn Asn Pro Val Leu Lys Phe Arg
35 40 45

Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala Phe Pro Ser
50 55 60

Leu Asn Thr Thr Thr Thr Pro Pro Ser Ile Gln Ser Leu Val Gln Glu
65 70 75 80

Phe Asp Pro Lys Ile Pro Ala Lys Asp Ala Leu Thr Pro Pro Ser Ser
85 90 95

Trp Tyr Thr Asp Ala Ala Phe Tyr Ala His Glu Leu Asp Arg Ile Phe
100 105 110

Tyr Lys Gly Trp Gln Val Pro Gly Tyr Ser Asp Gln Ile Lys Glu Pro
115 120 125

Asn Gln Tyr Phe Thr Gly Thr Leu Gly Asn Val Glu Tyr Leu Val Cys
130 135 140

Arg Asp Gly Glu Gly Lys Val His Ala Phe His Asn Val Cys Thr His
145 150 155 160

Arg Ala Ser Ile Leu Ala Cys Gly Ser Gly Lys Lys Ser Cys Phe Val
165 170 175

Cys Pro Tyr His Gly Trp Val Phe Gly Met Asn Gly Ser Leu Thr Lys
180 185 190

Ala Ser Lys Ala Thr Glu Glu Gln Ser Leu Asp Pro Asp Glu Leu Gly
195 200 205

Leu Val Pro Leu Lys Val Ala Val Trp Gly Pro Phe Ile Leu Ile Ser
210 215 220

Leu Asp Arg Ser Ser Leu Glu Val Gly Asp Val Gly Ser Glu Trp Leu
225 230 235 240

Gly Ser Cys Ala Glu Asp Val Lys Ala His Ala Phe Asp Pro Asn Leu
245 250 255

Gln Phe Ile Asn Arg Ser Glu Phe Pro Met Glu Ser Asn Trp Lys Ile
260 265 270

Phe Ser Asp Asn Tyr Leu Asp Ser Ser Tyr His Val Pro Tyr Ala His
275 280 285

Lys Tyr Tyr Ala Thr Glu Leu Asp Phe Asp Thr Tyr Gln Thr Asp Met
290 295 300

Ile Gly Asn Val Thr Ile Gln Arg Val Ala Gly Ser Ser Asn Asn Gly
305 310 315 320

Phe Asn Arg Leu Gly Ser Gln Ala Phe Tyr Ala Phe Ala Tyr Pro Asn
325 330 335

Phe Ala Val Glu Arg Tyr Gly Pro Trp Met Thr Thr Met His Ile Leu
340 345 350

Pro Leu Gly Pro Arg Lys Cys Lys Leu Val Val Asp Tyr Tyr Ile Glu
355 360 365

Lys Ser Lys Leu Asp Asp Lys Asp Tyr Ile Glu Lys Gly Ile Ala Ile
370 375 380

Asn Asp Asn Val Gln Lys Glu Asp Val Val Leu Cys Glu Ser Val Gln
385 390 395 400

Lys Gly Leu Glu Thr Pro Ala Tyr Arg Ser Gly Arg Tyr Val Met Pro
405 410 415

Ile Glu Lys Gly Ile His His Phe His Cys Trp Leu His Gln Val Leu
420 425 430

Lys

<210> 3

<211> 1651

<212> DNA

<213> Chenopodium album

<220>

<221> CDS

<222> (119)..(1423)

<223>

<400> 3

cttgaattat acaagctaag tatatatatt gttgatcatc tttcatacca cctttaaaaa 60

aaattataac aacaaaagga agtggttagt tattgcttga tcatcatata atatcaac 118

atg tca gca agt gca aca aca atg ttg ctg aaa tac cca aca act gta 166
Met Ser Ala Ser Ala Thr Thr Met Leu Leu Lys Tyr Pro Thr Thr Val
1 5 10 15

tgt ggt ata cca aat tca tca tca aac aat gat act tca aat aac atc 214
Cys Gly Ile Pro Asn Ser Ser Ser Asn Asn Asp Thr Ser Asn Asn Ile
20 25 30

gtc cca att cca caa act agt act aat aat ccg gta ctt aag ttt cgt 262
Val Pro Ile Pro Gln Thr Ser Thr Asn Asn Pro Val Leu Lys Phe Arg
35 40 45

acc cct aat aaa acc att aac gcc gtc gct gcc ccg gct ttt cct tct 310
Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala Phe Pro Ser
50 55 60

tta agt acc acc act act ccg ccg tcg att caa tca ctt gtc cag gaa 358
Leu Ser Thr Thr Thr Thr Pro Pro Ser Ile Gln Ser Leu Val Gln Glu
65 70 75 80

ttc gat ccg agg att ctg gcc gag gat gct ctc acg cct cct agc tct 406
Phe Asp Pro Arg Ile Leu Ala Glu Asp Ala Leu Thr Pro Pro Ser Ser
85 90 95

tgg tat act gaa cct gcc ttc tat gct cat gaa ctt gac cgt atc ttt 454
Trp Tyr Thr Glu Pro Ala Phe Tyr Ala His Glu Leu Asp Arg Ile Phe
100 105 110

tac	aaa	gga	tgg	caa	gtc	gca	ggg	tac	agc	gat	caa	att	aag	gag	cct	502
Tyr	Lys	Gly	Trp	Gln	Val	Ala	Gly	Tyr	Ser	Asp	Gln	Ile	Lys	Glu	Pro	
	115						120					125				
aac	caa	tat	ttc	acc	gga	acg	tta	gga	aat	gtt	gaa	tat	ttg	gtg	tgt	550
Asn	Gln	Tyr	Phe	Thr	Gly	Thr	Leu	Gly	Asn	Val	Glu	Tyr	Leu	Val	Cys	
	130						135				140					
cga	gat	ggt	gaa	gga	aaa	gtt	cat	gca	ttt	cac	aat	gtt	tgc	act	cat	598
Arg	Asp	Gly	Glu	Gly	Lys	Val	His	Ala	Phe	His	Asn	Val	Cys	Thr	His	
	145					150				155					160	
cgt	gct	tcg	att	ctt	gct	tgt	gga	agt	ggc	aaa	aaa	tcg	tgt	ttc	gta	646
Arg	Ala	Ser	Ile	Leu	Ala	Cys	Gly	Ser	Gly	Lys	Lys	Ser	Cys	Phe	Val	
				165					170					175		
tgc	cct	tac	cat	ggc	tgg	gta	ttt	ggc	atg	aat	gga	tca	ctt	acg	aaa	694
Cys	Pro	Tyr	His	Gly	Trp	Val	Phe	Gly	Met	Asn	Gly	Ser	Leu	Thr	Lys	
			180					185					190			
gct	tcc	aaa	gca	acc	gaa	gaa	cag	tcc	ctt	gat	ccc	gat	gaa	ctt	ggg	742
Ala	Ser	Lys	Ala	Thr	Glu	Glu	Gln	Ser	Leu	Asp	Pro	Asp	Glu	Leu	Gly	
		195					200					205				
ctt	gta	ccc	ctg	aaa	gtt	gca	gta	tgg	ggc	cca	ttt	ata	ctc	atc	agt	790
Leu	Val	Pro	Leu	Lys	Val	Ala	Val	Trp	Gly	Pro	Phe	Ile	Leu	Ile	Ser	
	210					215					220					
ttg	gac	aga	tca	agc	ctt	gaa	gta	ggc	gat	gtt	gga	tct	gaa	tgg	ctt	838
Leu	Asp	Arg	Ser	Ser	Leu	Glu	Val	Gly	Asp	Val	Gly	Ser	Glu	Trp	Leu	
	225				230					235					240	
ggc	agt	tgt	gct	gaa	gat	gtt	aag	gcc	cat	gct	ttt	gac	cct	aat	ttg	886
Gly	Ser	Cys	Ala	Glu	Asp	Val	Lys	Ala	His	Ala	Phe	Asp	Pro	Asn	Leu	
				245					250					255		
cag	ttc	atc	aat	agg	agt	gaa	ttt	cca	atg	gaa	tct	aat	tgg	aag	att	934
Gln	Phe	Ile	Asn	Arg	Ser	Glu	Phe	Pro	Met	Glu	Ser	Asn	Trp	Lys	Ile	
			260					265					270			
ttc	agt	gac	aac	tac	ttg	gat	agc	tcg	tac	cat	gtt	cct	tat	gca	cac	982
Phe	Ser	Asp	Asn	Tyr	Leu	Asp	Ser	Ser	Tyr	His	Val	Pro	Tyr	Ala	His	
		275					280					285				
aag	tac	tat	gca	act	gaa	ctc	gac	ttt	gat	act	tat	caa	acc	gat	atg	1030
Lys	Tyr	Tyr	Ala	Thr	Glu	Leu	Asp	Phe	Asp	Thr	Tyr	Gln	Thr	Asp	Met	
	290					295					300					
att	gga	aat	gtc	acg	att	caa	aga	gtg	gcg	ggg	agt	tca	aac	aag	cca	1078
Ile	Gly	Asn	Val	Thr	Ile	Gln	Arg	Val	Ala	Gly	Ser	Ser	Asn	Lys	Pro	
	305				310					315					320	

gat ggt ttt gat aga ctt gga tct caa gca ttc tat gct ttt gca tac	1126
Asp Gly Phe Asp Arg Leu Gly Ser Gln Ala Phe Tyr Ala Phe Ala Tyr	
325 330 335	
cct aac ttt gct gtg gaa agg tat ggc cct tgg atg aca aca atg cat	1174
Pro Asn Phe Ala Val Glu Arg Tyr Gly Pro Trp Met Thr Thr Met His	
340 345 350	
att ctt cca tta gga cca aga aaa tgc aaa tta gtg gtg gac tac tat	1222
Ile Leu Pro Leu Gly Pro Arg Lys Cys Lys Leu Val Val Asp Tyr Tyr	
355 360 365	
att gaa aaa tca atg ctg gac gac aag gat tac atc gag aag ggc ata	1270
Ile Glu Lys Ser Met Leu Asp Asp Lys Asp Tyr Ile Glu Lys Gly Ile	
370 375 380	
gca atc aat gat aat gta cag aaa gaa gat gtg gtg ttg tgt gaa agt	1318
Ala Ile Asn Asp Asn Val Gln Lys Glu Asp Val Val Leu Cys Glu Ser	
385 390 395 400	
gtc caa aaa ggg ttg gag aca cca gca tat cgt agt gga aga tat gtg	1366
Val Gln Lys Gly Leu Glu Thr Pro Ala Tyr Arg Ser Gly Arg Tyr Val	
405 410 415	
atg cca att gag aaa gga atc cat cat ttc cac tgt tgg ttg cac caa	1414
Met Pro Ile Glu Lys Gly Ile His His Phe His Cys Trp Leu His Gln	
420 425 430	
gta ttg aag tgatagcagc agatcagatg ttcgtttcctt aatttccttt	1463
Val Leu Lys	
435	
tattggaact ggataattat aataataata agtaaaaaag taaaattata atgtcatgta	1523
gttgagattg ttgctagagt tgagcgtatg ctccatcatgc acttagttat caagtgtgta	1583
tgtgttttgt catggacaaa atgtttccttg ctagaattta tcatattata aggtgcta	1643
gtccaata	1651

<210> 4

<211> 435

<212> PRT

<213> Chenopodium album

<400> 4

Met	Ser	Ala	Ser	Ala	Thr	Thr	Met	Leu	Leu	Lys	Tyr	Pro	Thr	Thr	Val	1	5	10	15
Cys	Gly	Ile	Pro	Asn	Ser	Ser	Ser	Asn	Asn	Asp	Thr	Ser	Asn	Asn	Ile	20	25	30	
Val	Pro	Ile	Pro	Gln	Thr	Ser	Thr	Asn	Asn	Pro	Val	Leu	Lys	Phe	Arg	35	40	45	
Thr	Pro	Asn	Lys	Thr	Ile	Asn	Ala	Val	Ala	Ala	Pro	Ala	Phe	Pro	Ser	50	55	60	
Leu	Ser	Thr	Thr	Thr	Thr	Pro	Pro	Ser	Ile	Gln	Ser	Leu	Val	Gln	Glu	65	70	75	80
Phe	Asp	Pro	Arg	Ile	Leu	Ala	Glu	Asp	Ala	Leu	Thr	Pro	Pro	Ser	Ser	85	90	95	
Trp	Tyr	Thr	Glu	Pro	Ala	Phe	Tyr	Ala	His	Glu	Leu	Asp	Arg	Ile	Phe	100	105	110	
Tyr	Lys	Gly	Trp	Gln	Val	Ala	Gly	Tyr	Ser	Asp	Gln	Ile	Lys	Glu	Pro	115	120	125	
Asn	Gln	Tyr	Phe	Thr	Gly	Thr	Leu	Gly	Asn	Val	Glu	Tyr	Leu	Val	Cys	130	135	140	
Arg	Asp	Gly	Glu	Gly	Lys	Val	His	Ala	Phe	His	Asn	Val	Cys	Thr	His	145	150	155	160
Arg	Ala	Ser	Ile	Leu	Ala	Cys	Gly	Ser	Gly	Lys	Lys	Ser	Cys	Phe	Val	165	170	175	
Cys	Pro	Tyr	His	Gly	Trp	Val	Phe	Gly	Met	Asn	Gly	Ser	Leu	Thr	Lys	180	185	190	
Ala	Ser	Lys	Ala	Thr	Glu	Glu	Gln	Ser	Leu	Asp	Pro	Asp	Glu	Leu	Gly	195	200	205	

Leu Val Pro Leu Lys Val Ala Val Trp Gly Pro Phe Ile Leu Ile Ser
210 215 220

Leu Asp Arg Ser Ser Leu Glu Val Gly Asp Val Gly Ser Glu Trp Leu
225 230 235 240

Gly Ser Cys Ala Glu Asp Val Lys Ala His Ala Phe Asp Pro Asn Leu
245 250 255

Gln Phe Ile Asn Arg Ser Glu Phe Pro Met Glu Ser Asn Trp Lys Ile
260 265 270

Phe Ser Asp Asn Tyr Leu Asp Ser Ser Tyr His Val Pro Tyr Ala His
275 280 285

Lys Tyr Tyr Ala Thr Glu Leu Asp Phe Asp Thr Tyr Gln Thr Asp Met
290 295 300

Ile Gly Asn Val Thr Ile Gln Arg Val Ala Gly Ser Ser Asn Lys Pro
305 310 315 320

Asp Gly Phe Asp Arg Leu Gly Ser Gln Ala Phe Tyr Ala Phe Ala Tyr
325 330 335

Pro Asn Phe Ala Val Glu Arg Tyr Gly Pro Trp Met Thr Thr Met His
340 345 350

Ile Leu Pro Leu Gly Pro Arg Lys Cys Lys Leu Val Val Asp Tyr Tyr
355 360 365

Ile Glu Lys Ser Met Leu Asp Asp Lys Asp Tyr Ile Glu Lys Gly Ile
370 375 380

Ala Ile Asn Asp Asn Val Gln Lys Glu Asp Val Val Leu Cys Glu Ser
385 390 395 400

Val Gln Lys Gly Leu Glu Thr Pro Ala Tyr Arg Ser Gly Arg Tyr Val
405 410 415

Met Pro Ile Glu Lys Gly Ile His His Phe His Cys Trp Leu His Gln
 420 425 430

Val Leu Lys
 435

<210> 5

<211> 1712

<212> DNA

<213> Chenopodium album

<220>

<221> CDS

<222> (133)..(1431)

<223>

<400> 5

cttgaattac acaagctaag ttaagctaag ctatattggt gatcatcttt cataccacct 60

cctttaaaaa aaaaaaatta taacaacaaa aggaagtgtt tagttattgc ttgatcatca 120

tataacatca at atg gca gca agt gca aca aca atg ttg ctg aaa tac cca 171
 Met Ala Ala Ser Ala Thr Thr Met Leu Leu Lys Tyr Pro
 1 5 10

aca act gta tgt ggt ata cca aat tca tca tca aac aat gat act tca 219
 Thr Thr Val Cys Gly Ile Pro Asn Ser Ser Ser Asn Asn Asp Thr Ser
 15 20 25

aat aac atc gtc cca att cca caa act att act aat aat ccg gta ctt 267
 Asn Asn Ile Val Pro Ile Pro Gln Thr Ile Thr Asn Asn Pro Val Leu
 30 35 40 45

aag ttt cgt acc cct aat aaa acc att aac gcc gtc gct gcc ccg gct 315
 Lys Phe Arg Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala
 50 55 60

ttt cct tct tta aac acc acc act act ccg ccg tca att caa tca ctt 363
 Phe Pro Ser Leu Asn Thr Thr Thr Thr Pro Pro Ser Ile Gln Ser Leu
 65 70 75

gtc	cag	gaa	ttc	gat	ccg	agg	att	ccg	gcc	gag	gat	gct	ctt	acg	cct	411
Val	Gln	Glu	Phe	Asp	Pro	Arg	Ile	Pro	Ala	Glu	Asp	Ala	Leu	Thr	Pro	
	80						85					90				
cct	agc	tct	tgg	tat	act	gaa	cct	gct	ttc	tat	gct	cat	gaa	ctt	gac	459
Pro	Ser	Ser	Trp	Tyr	Thr	Glu	Pro	Ala	Phe	Tyr	Ala	His	Glu	Leu	Asp	
	95					100					105					
cgt	atc	ttt	tac	aag	gga	tgg	caa	gtc	gca	ggg	tac	agt	gat	caa	att	507
Arg	Ile	Phe	Tyr	Lys	Gly	Trp	Gln	Val	Ala	Gly	Tyr	Ser	Asp	Gln	Ile	
110					115					120					125	
aag	gag	cct	aac	caa	tat	ttc	acc	gga	acg	tta	gga	aat	gtt	gaa	tat	555
Lys	Glu	Pro	Asn	Gln	Tyr	Phe	Thr	Gly	Thr	Leu	Gly	Asn	Val	Glu	Tyr	
				130					135					140		
ttg	gtg	tgt	cga	gat	ggg	gaa	ggg	aaa	gtt	cat	gca	ttt	cac	aac	gtt	603
Leu	Val	Cys	Arg	Asp	Gly	Glu	Gly	Lys	Val	His	Ala	Phe	His	Asn	Val	
			145					150					155			
tgc	acc	cat	cgt	gct	tcg	att	ctt	gct	tgt	gga	agc	gga	aaa	aaa	tcg	651
Cys	Thr	His	Arg	Ala	Ser	Ile	Leu	Ala	Cys	Gly	Ser	Gly	Lys	Lys	Ser	
		160					165					170				
tgt	ttt	gta	tgc	cct	tac	cat	gga	tgg	gta	ttt	ggc	atg	aat	gga	tcg	699
Cys	Phe	Val	Cys	Pro	Tyr	His	Gly	Trp	Val	Phe	Gly	Met	Asn	Gly	Ser	
	175					180					185					
ctt	aca	aaa	gct	tcc	aaa	gca	agc	gaa	gaa	cag	tca	ctt	gat	ccc	gat	747
Leu	Thr	Lys	Ala	Ser	Lys	Ala	Ser	Glu	Glu	Gln	Ser	Leu	Asp	Pro	Asp	
190					195					200					205	
gaa	ctt	ggg	ctt	gta	ccc	ctg	aaa	gtt	gca	gta	tgg	ggc	cca	ttt	ata	795
Glu	Leu	Gly	Leu	Val	Pro	Leu	Lys	Val	Ala	Val	Trp	Gly	Pro	Phe	Ile	
				210					215					220		
ctc	atc	agt	ttg	gac	aga	tca	agc	ctt	gaa	gta	gat	gat	gtt	gga	tct	843
Leu	Ile	Ser	Leu	Asp	Arg	Ser	Ser	Leu	Glu	Val	Asp	Asp	Val	Gly	Ser	
			225					230					235			
gaa	tgg	ctt	ggg	agt	tgt	gct	gaa	gat	gtt	aag	gcc	cat	gct	ttt	gac	891
Glu	Trp	Leu	Gly	Ser	Cys	Ala	Glu	Asp	Val	Lys	Ala	His	Ala	Phe	Asp	
		240					245					250				
cct	aat	ttg	cag	ttc	atc	aat	agg	agt	gaa	ttt	cca	atg	gaa	tct	aat	939
Pro	Asn	Leu	Gln	Phe	Ile	Asn	Arg	Ser	Glu	Phe	Pro	Met	Glu	Ser	Asn	
	255					260					265					
tgg	aag	att	ttc	agt	gac	aac	tat	ttg	gat	agc	tcg	tac	cat	gtt	cct	987
Trp	Lys	Ile	Phe	Ser	Asp	Asn	Tyr	Leu	Asp	Ser	Ser	Tyr	His	Val	Pro	
270					275					280					285	

tat gca cac aag tac tat gct act gaa ctc gac ttt gat act tac caa	1035
Tyr Ala His Lys Tyr Tyr Ala Thr Glu Leu Asp Phe Asp Thr Tyr Gln	
290 295 300	
act gat atg atc gga aat gtc acg att caa aga gtg gca ggg agt tca	1083
Thr Asp Met Ile Gly Asn Val Thr Ile Gln Arg Val Ala Gly Ser Ser	
305 310 315	
aac aat ggt ttt aat aga ctt gga tct caa gca ttc tac gct ttt gca	1131
Asn Asn Gly Phe Asn Arg Leu Gly Ser Gln Ala Phe Tyr Ala Phe Ala	
320 325 330	
tac cct aac ttt gct gtg gaa agg tat ggc cct tgg atg aca aca atg	1179
Tyr Pro Asn Phe Ala Val Glu Arg Tyr Gly Pro Trp Met Thr Thr Met	
335 340 345	
cac att ctt cca tta gga cca agg aaa tgc aaa tta gtg gtg gac tac	1227
His Ile Leu Pro Leu Gly Pro Arg Lys Cys Lys Leu Val Val Asp Tyr	
350 355 360 365	
tat att gaa aaa tca aag ctg gac gac aag gat tac atc gag aag ggc	1275
Tyr Ile Glu Lys Ser Lys Leu Asp Asp Lys Asp Tyr Ile Glu Lys Gly	
370 375 380	
ata gca atc aat gat aat gta cag aaa gaa gat gtg gtg ttg tgt gaa	1323
Ile Ala Ile Asn Asp Asn Val Gln Lys Glu Asp Val Val Leu Cys Glu	
385 390 395	
agt gtc caa aaa ggg ttg gag aca cct gcg tat cgt agt gga aga tat	1371
Ser Val Gln Lys Gly Leu Glu Thr Pro Ala Tyr Arg Ser Gly Arg Tyr	
400 405 410	
gtg atg cca att gag aaa gga atc cat cat ttc cac tgt tgg ttg cac	1419
Val Met Pro Ile Glu Lys Gly Ile His His Phe His Cys Trp Leu His	
415 420 425	
caa gta ttg aag tgattgcagc agatcagatg ttcggtttcctt aatttccttt	1471
Gln Val Leu Lys	
430	
tattggaatt ggatgattgt tataataata agtaaaatta taatgtcatg tagttgagat	1531
tgttgctaga gttgagcgta tgctcctcat gcacttagtt atcaagtgtg tatgtgtttg	1591
gtcatgggca aaatgtatatt tcttgctaga atttggttata ttatgggtgct aatgtccaat	1651
aatataaata acaccattgc accctttccc tacttgagaa attatatccc atttattttc	1711
g	1712

<210> 6

<211> 433

<212> PRT

<213> Chenopodium album

<400> 6

Met Ala Ala Ser Ala Thr Thr Met Leu Leu Lys Tyr Pro Thr Thr Val
1 5 10 15

Cys Gly Ile Pro Asn Ser Ser Ser Asn Asn Asp Thr Ser Asn Asn Ile
20 25 30

Val Pro Ile Pro Gln Thr Ile Thr Asn Asn Pro Val Leu Lys Phe Arg
35 40 45

Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala Phe Pro Ser
50 55 60

Leu Asn Thr Thr Thr Thr Pro Pro Ser Ile Gln Ser Leu Val Gln Glu
65 70 75 80

Phe Asp Pro Arg Ile Pro Ala Glu Asp Ala Leu Thr Pro Pro Ser Ser
85 90 95

Trp Tyr Thr Glu Pro Ala Phe Tyr Ala His Glu Leu Asp Arg Ile Phe
100 105 110

Tyr Lys Gly Trp Gln Val Ala Gly Tyr Ser Asp Gln Ile Lys Glu Pro
115 120 125

Asn Gln Tyr Phe Thr Gly Thr Leu Gly Asn Val Glu Tyr Leu Val Cys
130 135 140

Arg Asp Gly Glu Gly Lys Val His Ala Phe His Asn Val Cys Thr His
145 150 155 160

Arg Ala Ser Ile Leu Ala Cys Gly Ser Gly Lys Lys Ser Cys Phe Val
165 170 175

Cys Pro Tyr His Gly Trp Val Phe Gly Met Asn Gly Ser Leu Thr Lys
180 185 190

Ala Ser Lys Ala Ser Glu Glu Gln Ser Leu Asp Pro Asp Glu Leu Gly
195 200 205

Leu Val Pro Leu Lys Val Ala Val Trp Gly Pro Phe Ile Leu Ile Ser
210 215 220

Leu Asp Arg Ser Ser Leu Glu Val Asp Asp Val Gly Ser Glu Trp Leu
225 230 235 240

Gly Ser Cys Ala Glu Asp Val Lys Ala His Ala Phe Asp Pro Asn Leu
245 250 255

Gln Phe Ile Asn Arg Ser Glu Phe Pro Met Glu Ser Asn Trp Lys Ile
260 265 270

Phe Ser Asp Asn Tyr Leu Asp Ser Ser Tyr His Val Pro Tyr Ala His
275 280 285

Lys Tyr Tyr Ala Thr Glu Leu Asp Phe Asp Thr Tyr Gln Thr Asp Met
290 295 300

Ile Gly Asn Val Thr Ile Gln Arg Val Ala Gly Ser Ser Asn Asn Gly
305 310 315 320

Phe Asn Arg Leu Gly Ser Gln Ala Phe Tyr Ala Phe Ala Tyr Pro Asn
325 330 335

Phe Ala Val Glu Arg Tyr Gly Pro Trp Met Thr Thr Met His Ile Leu
340 345 350

Pro Leu Gly Pro Arg Lys Cys Lys Leu Val Val Asp Tyr Tyr Ile Glu
355 360 365

Lys Ser Lys Leu Asp Asp Lys Asp Tyr Ile Glu Lys Gly Ile Ala Ile
370 375 380

Asn Asp Asn Val Gln Lys Glu Asp Val Val Leu Cys Glu Ser Val Gln
385 390 395 400

Lys Gly Leu Glu Thr Pro Ala Tyr Arg Ser Gly Arg Tyr Val Met Pro
405 410 415

Ile Glu Lys Gly Ile His His Phe His Cys Trp Leu His Gln Val Leu
420 425 430

Lys

<210> 7

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> misc_feature

<222> (9)..(9)

<223> n is a, g, c, or t

<220>

<221> misc_feature

<222> (15)..(15)

<223> n is a, g, c, or t

<220>

<221> misc_feature

<222> (18)..(18)

<223> n is a, g, c, or t

<400> 7

tggtayacng arccngcntt yta

23

<210> 8

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> misc_feature

<222> (9)..(9)

<223> n is a, g, c, or t

<220>

<221> misc_feature

<222> (15)..(15)

<223> n is a, g, c, or t

<220>

<221> misc_feature

<222> (18)..(18)

<223> n is a, g, c, or t

<400> 8
tayttrtgng crtanggnac rtgrta

26

<210> 9

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 9
gtgcattggt gtcaccaag ggcc

24

<210> 10

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 10
gatcccgatg aacttgggct tgtacccc

28

<210> 11

<211> 26

<212> DNA

<213> Artificial Sequence

<220>
<223> Synthetic DNA
<400> 11
cccgggttta gttattgctt gatcat 26

<210> 12
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic DNA
<400> 12
gagctcctgc aatcacttca atactt 26

<210> 13
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic DNA
<400> 13
taatggatcc attaacgccg tcgc 24

<210> 14
<211> 25
<212> DNA
<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 14

gggtaccaat cacttcaata cttgg

25

<210> 15

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 15

cccgggaaaa ccattatggc cgtcgc

26

<210> 16

<211> 127

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> CDS

<222> (1)..(126)

<223>

<400> 16

atg	gca	gca	agt	gca	aca	aca	atg	ttg	ctg	aaa	tac	cca	aca	act	gta	48
Met	Ala	Ala	Ser	Ala	Thr	Thr	Met	Leu	Leu	Lys	Tyr	Pro	Thr	Thr	Val	
1				5				10						15		

tgt	ggt	ata	cca	aat	tca	tca	tca	aac	aat	gat	act	tca	aat	aac	atc	96
Cys	Gly	Ile	Pro	Asn	Ser	Ser	Ser	Asn	Asn	Asp	Thr	Ser	Asn	Asn	Ile	
			20					25					30			

gtc	cca	att	cca	caa	act	att	act	aat	aat	c						127
Val	Pro	Ile	Pro	Gln	Thr	Ile	Thr	Asn	Asn							
		35					40									

<210> 17

<211> 42

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 17

Met	Ala	Ala	Ser	Ala	Thr	Thr	Met	Leu	Leu	Lys	Tyr	Pro	Thr	Thr	Val
1				5				10						15	

Cys	Gly	Ile	Pro	Asn	Ser	Ser	Ser	Asn	Asn	Asp	Thr	Ser	Asn	Asn	Ile
			20					25					30		

AI Cons

Val	Pro	Ile	Pro	Gln	Thr	Ile	Thr	Asn	Asn						
		35					40								